HIGHER EDUCATION COORDINATING COMMISSION
255 Capitol Street NE, Salem, OR 97310
www.oregon.gov/HigherEd

Docket Item:

University Program Approval: University of Oregon, Bachelor of Arts (B.A.)/Bachelor of Science (B.S.) in Data Science

Summary:

University of Oregon proposes a new degree program leading to a B.A./B.S. in Data Science. The statewide Provosts' Council has unanimously recommended approval. Higher Education Coordinating Commission (HECC) staff completed a review of the proposed program. After analysis, HECC staff recommends approval of the program as proposed.

Staff Recommendation:

The HECC recommends the adoption of the following resolution:

> RESOLVED, that the Higher Education Coordinating Commission approve the following program: B.A./B.S. in Data Science at the University of Oregon

HIGHER EDUCATION COORDINATING COMMISSION

Proposal for a New Academic Program

Institution: University of Oregon

College/School: College of Arts & Sciences (CAS)

Department/Program Name: Data Science Program

Degree and Program Title: Bachelor of Arts and Bachelor of Science in Data Science

1. Program Description

a. Proposed Classification of Instructional Programs (CIP) number.

30.7001

b. Brief overview (1-2 paragraphs) of the proposed program, including its disciplinary foundations and connections; program objectives; programmatic focus; degree, certificate, minor, and concentrations offered.

We propose an undergraduate degree in data science that is quantitatively focused, requiring courses in computer science, mathematics, probability and statistics, and machine learning. This quantitative core of courses integrates with applications of these concepts and tools in specific areas such as biology, finance, geography, linguistics, etc. (so called "domain" areas). The degree finishes with the option of a capstone project demonstrating mastery of data science concepts within the student's chosen domain of specialization.

Following examination of successful programs at other top universities, we closely followed the curriculum in the Division of Data Science and Information at UC Berkeley. This program is widely recognized as a model, and has explicitly developed data science teaching material to be utilized by other programs and universities. We have adopted and modified proposed entry and mid-level courses in core data science proficiencies and exposure to domain emphases, as well as steps through key offerings in math, computer science, ethics, and communication. The core data science training is aligned with existing courses applying data science techniques in several units at UO.

c. Course of study – proposed curriculum, including course numbers, titles, and credit hours.

The data science curriculum combines general principles with domain-specific application. The curriculum is sub-divided into the following categories with the corresponding requirements:

Core data science courses (16 credits):

- DSCI 101 – Foundations of Data Science I
- DSCI 102 – Foundations of Data Science II
- DSCI 311 – Principles and Techniques of Data Science

- DSCI 411 – Data Science Capstone Project*

*An additional domain-specialization course may be substituted in lieu of DSCI 411.

Foundations in Mathematics and Computing (28 credits):

- Foundational computer science series: CIS 210, 211, 212
- Foundational Calculus series: MATH 251, 252
- Linear Algebra Series: MATH 341-342

Computational and Inferential Depth (12 credits)

Students must select **three** courses from the following list below. Please see the note at the end of this section regarding Computer and Information Sciences (CIS) prerequisites in the Computational and Inferential Depth category.

- CIS 313 – Intermediate Data Structures
- CIS 314 – Computer Organization
- CIS 322 – Introduction to Software Engineering
- CIS 333 – Applied Cryptography
- CIS 451 – Intro to Databases
- CIS 315 – Intermediate Algorithms
- CIS 330 – C/C++ & Unix
- CIS 410vis – Visualization
- CIS 415 – Operating Systems
- CIS 432 – Intro to Internet
- CIS 434 – Distributed Systems and Network Security
- CIS 425 – Programming Languages
- CIS 422 – Software Engineering
- CIS 471 – Intro to AI

Modeling, Learning and Decision Making (4 credits)

- DSCI/CIS 372 – Machine Learning for Data Science

Probability (4 credits)

- DSCI/MATH 345 – Probability and Statistics for Data Science

Human Contexts and Ethics (4 credits)

- PHIL 423 – Technology Ethics

Domain Emphasis (20-24 credits)

The domain emphasis consists of completing 2-3 courses (8-12 credits) in the domain core, followed by a minimum of 3 courses (12 credits) of domain specialization. For each domain emphasis, a curated list of courses has been developed for both the core and specialization component. See example below.

Linguistics Domain:
Domain Core – take both

- LING 301 – Linguistic Analysis 4 cr
- LING 302 – Linguistic Behavior 4 cr

Domain Specialization – select at least three

- LING 435 – Morphology & Syntax 4cr
- LING 451 – Functional Syntax I 4cr
- LING 452 – Functional Syntax II 4cr
- LING 493 – Corpus Linguistics 4cr
- CIS 410nlp – Natural Language Processing 4cr

d. Manner in which the program will be delivered, including program location (if offered outside of the main campus), course scheduling, and the use of technology (for both on-campus and off-campus delivery).

The BA/BS in Data Science will be offered at the University of Oregon home campus in Eugene. We are not currently planning to offer the degree online but some of the courses may be offered online.

e. Adequacy and quality of faculty delivering the program.

Because the major is meant to provide foundational training that will be applied through applications to domains in partnership with existing units (see program description), the development of the program primarily involves the inclusion of existing courses and faculty at UO. For example, core courses in math and CIS in applied mathematics, statistics, and computer science either already exist or were planned to be offered by faculty in those units in collaboration with the Data Science Program.

In addition, over the past several years faculty lines have been approved as being affiliated with the UO Presidential Data Science Initiative (DSI), launched by President Schill in 2017. These faculty members are planning courses that will contribute to the core offerings (e.g., data science ethics will be offered via faculty hired into philosophy in CAS) and domain offerings at the upper division (e.g., hires in biology, psychology, business, education, etc.).

As outlined in the program description, only a handful of new courses will need to be developed (e.g. DSCI 101, 102, 311, 345, 372, 402) specifically to fulfill the pedagogical needs of training in core data science principles. These courses will be developed through two primary mechanisms. First, working under the guidance of TTF members of the Data Science Academic Leadership Team, key NTTF in DSI will work to build the courses and begin delivering content. These individuals are highly trained in data science and have been developing and teaching courses in a variety of data science areas at UO. Second, key interested TTF faculty, several of whom were hired in association with DSI, will be recruited to participate in the development and delivery of these courses via course buyouts to home

departments. Lastly, as these courses and the program grows, additional faculty will be added. At maturity, the suite of core data science courses will be offered by core faculty in the DSI.

Leadership of the Data Science Program will continue to occur initially (first two years) through the DSI, with Deputy Director Joe Sventek continuing to play the role of program director. Coordination of the program with the larger research and graduate educational missions of the DSI will occur under the guidance of DSI's Executive Director Bill Cresko. As the program grows, the addition of support staff and leadership will occur organically via existing allocation processes at the university. The anticipation is that at maturity - approximately 4 to 5 years - the program will have the full leadership and support staff typical of a large department in CAS.

f. Adequacy of faculty resources – full-time, part-time, adjunct.

See above

g. Other staff.

FTE will be built into the budget through staff and peer mentors to provide additional, data science-specific advising support for students majoring in data science. These resources will be appropriately scaled as the program grows in accordance with existing allocation practices. We anticipate that at full scale, a 1.0 FTE advising role will be provided.

Full scale anticipated staffing needs:
- 1.0 FTE Department/Office Manager
- 1.0 FTE Undergraduate Administrator
- .5 FTE Director of Undergraduate Studies
- 1.0 FTE Advisor
- 33 GE Terms
- 1.0 Capstone and Internship Coordinator (3 years out)
- 1.0 Technology and Lab preparator
- Program Director – Service assignment for TTF

h. Adequacy of facilities, library, and other resources.

Computational infrastructure exists to support this degree due to the UO's commitment to increasing IT infrastructure over the previous 5 years. In particular, the creation of the high performance computing structure Talapas have made UO's computational capabilities world class. We will require specialized local equipment, such as a laptop pool, an educational condo on Talapas, and support for the Jupyter collaborative coding environment.

In addition to equipment, we will require large classroom space for lectures and spaces for interactive computer education. The latter types of spaces will be our main constraint over the short term. The collaborative rooms in the Price Science Library (B042 and B044) are exemplars for this type of 'flipped' classroom but are in high demand. However, the university plans to add such dry interactive spaces throughout the next decade. For example, the opening of the Phil and Penny Knight Campus for Accelerating Scientific Impact will provide the possibility for sharing six of these types of rooms.

We anticipate no unusual library needs or other extended infrastructure resources at this time.

i. Anticipated start date.
Fall 2020

2. **Relationship to Mission and Goals**

   a. **Manner in which the proposed program supports the institution's mission, signature areas of focus, and strategic priorities.**

   The University of Oregon is a comprehensive public research university committed to exceptional teaching, discovery, and service. Because data science is a growing interdisciplinary field, a data science undergraduate degree program is essential to UO's mission and is part of a key presidential initiative. The collaborative nature of this degree builds upon UO strengths in interdisciplinary programs.

   b. **Manner in which the proposed program contributes to institutional and statewide goals for student access and diversity, quality learning, research, knowledge creation and innovation, and economic and cultural support of Oregon and its communities.**

   Accessibility, equity and inclusion are fundamental to the data science program. By nature, data science involves and impacts all members of society. Correspondingly, an essential measure of success for the program is the recruitment and sustained support of a richly diverse student cohort, particularly for students traditionally underrepresented in STEM fields. We plan to intentionally include advising and activities to attract and support such a diverse community of students and scholars.

   Specific efforts with respect to education and quality learning involve adopting a pedagogical approach that embraces students who may not otherwise be exposed to STEM fields or who may be intimidated by quantitative STEM programs. The first foundational data science course, DSCI 101, has been developed to minimize barriers to entry. It has no prerequisites and embraces students with no prior computing or statistical background. It will be offered as a UO core education course with the intent to have as broad of a reach and impact as possible. In-class demonstrations and examples are intentionally sourced to include culturally inclusive content and to address a variety of learning styles. As students progress through the lower- and upper-division DSCI curriculum, ethics and human context are woven throughout each of the DSCI courses.

   Further initiatives will emulate the successful practices employed by CIS, including active recruitment of women and minorities into the program and summer internship placements with companies partnering with the data science program. It is expected that these measures will ensure successful employment of data science students, including those traditionally underrepresented, as has been the case for CIS.

   c. **Manner in which the program meets regional or statewide needs and enhances the state's capacity to:**

      i. **improve educational attainment in the region and state;**
      ii. **respond effectively to social, economic, and environmental challenges and opportunities; and**
      iii. **address civic and cultural demands of citizenship.**

   With so many different domains requiring data science expertise, the core + domain emphasis built into the proposed program will help students to develop the core quantitative skills and apply them immediately to their domain of interest. We expect graduates of this program to find quality jobs immediately after graduation, both in the local area and throughout Oregon and the rest of the US.

In alignment with the deep history of UO as a liberal arts research university, a goal of the program is to not only help students utilize data science but to think critically about the impacts of data science on society. The program requires one course in data ethics, with other ethics courses available as area electives; three data science core courses (DSCI 101, 102, 311) also include ethical topics in their syllabi; finally, courses in cybersecurity are also available as area electives. This is an important aspect of the program addressing civic and cultural demands of citizenship - securing and ethically using domain data is essential to prevent misuse of said data.

3. Accreditation

Data science is a new discipline and as such there is no accreditation standard or organization that accredits undergraduate data science programs. Once such an accreditation organization and standard has been developed, we will reconsider accrediting the program.

4. Need

a. Anticipated fall term headcount and FTE enrollment over each of the next five years.

Given the market demand and career opportunities for Data Science, we expect this to be a popular degree option for UO students, and that we will attract students who may not have considered coming to UO prior to this degree option. As such, we've estimated potential enrollments accordingly.

Year 1	Year 2	Year 3	Year 4	Year 5
25 enrolled	100 incoming 125 enrolled	150 incoming 250 enrolled	200 Incoming 405 enrolled	200 incoming 490 enrolled

b. Expected degrees/certificates produced over the next five years.

Year 1	Year 2	Year 3	Year 4	Year 5
0	0	0	25	100

c. Characteristics of students to be served (resident/nonresident/international; traditional/ nontraditional; full-time/part-time, etc.).

Students pursing the BA/BS in Data Science will be resident, nonresident, and international full-time traditional students.

d. Evidence of market demand.

Experienced data scientists are a growing economic need. The January 2019 report from Indeed.com, one of the top job websites, showed a 29% increase in demand for data scientists year over year, and a 344% increase since 2013. Data from the technology job website Dice.com showed a 32% increase in data science job postings year over year.

Dice.com also noted that the job postings are from companies in a wide variety of industries, not just technology – e.g., investment banking, insurance, healthcare. Indeed.com currently lists 89 open data science jobs in Oregon; if the current year-on-year growth rate continues,

this means that there will be 317 open data science jobs on this date in 2024 when we produce our first group of graduates.

Today, Indeed.com shows over 10,000 open data science jobs nationally. With such growth in the demand for trained data scientists in Oregon and across the country, training in data science will become ubiquitous across academia.

e. If the program's location is shared with another similar Oregon public university program, the proposal should provide externally validated evidence of need (e.g., surveys, focus groups, documented requests, occupational/employment statistics and forecasts).

N/A

f. Estimate the prospects for success of program graduates (employment or graduate school) and consideration of licensure, if appropriate. What are the expected career paths for students in this program?

Data science is a growing field with increasing employment demands. Graduates of the UO's data science program will be competitive for data scientist positions in industry, including data analyst, quantitative analyst, data engineer, or artificial intelligence/machine learning engineer. The demand for these positions has grown substantially in all sectors with the increasing availability of business and customer data. The workforce analysis company Glassdoor.com estimates starting salaries between $81k-$149k for entry-level data scientists.

UO's program requires experience in a domain area which would offer additional employment opportunities. A domain emphasis in biology, for example, would prepare graduates for careers in bioinformatics or biostatistics. Qualified students will have the option to complete a capstone project to demonstrate competency to potential employers, which would increase their competitiveness in the job market.

5. Outcomes and Quality Assessment

a. Expected learning outcomes of the program.

Principle Learning Outcome (Concept or Skill)	Part of curriculum where this is introduced	Part of curriculum where this is developed	Where students demonstrate mastery
Have demonstrated the ability to assess data set quality, identifying and rectifying potential errors in such a way so as to lead to statistically meaningful derived information	DSCI 101	DSCI 102	DSCI 311 DSCI 411
Be able to visualize complex data sets using descriptive statistics and graphs	DSCI 101	DSCI 102	DSCI 311 DSCI 411
Demonstrate understanding of basic regression, optimization, prediction, simulation, and visualization methods	DSCI 102	DSCI 311	DSCI 345

Principle Learning Outcome (Concept or Skill)	Part of curriculum where this is introduced	Part of curriculum where this is developed	Where students demonstrate mastery
Be able to use critical thinking skills to translate substantive questions into well-defined statistical or probability problems and choose the appropriate graphical or numerical descriptive and/or inferential statistical techniques for a given problem, leading to actionable, valid, and meaningful conclusions	DSCI 311	DSCI 345	DSCI 411
Have developed successful strategies for formulating and testing hypotheses about data	DSCI 311	DSCI 345	DSCI 411
Have demonstrated an understanding of ethical, legal, societal, and economic concerns	DSCI 101 DSCI 102	DSCI 311	DSCI 411 PHIL 423
Be able to apply fundamental concepts of data science (data management, statistical prediction and inference, experimental design, etc.) to applications specific to the chosen specialization domain			DSCI 411

b. Methods by which the learning outcomes will be assessed and used to improve curriculum and instruction.

The assessment of degree learning outcomes will be derived from a sub-sample of artefacts generated through the program. The following will enable us to track each outcome from cohort to cohort:

1. Mastery of this learning outcome is covered in DSCI 311; one project that students must submit for assessment in 311 is to take a dirty data set, representing all the types of imperfections covered in the courses, and using techniques that they have learned in 101, 102, and 311, they must clean the data set for a standard statistical analysis.

2. One project that students must submit for assessment in 102 is to compute standard statistical attributes for a given data set and to use an appropriate graphical representation to communicate those attributes.

3. This will be assessed through one of the assessment mechanisms in DSCI 345.

4. This will be assessed through a detailed analysis of the DSCI 411 project reports, particularly establishing an hypothesis of information buried in the data, applying suitable

and appropriate analysis techniques to the data, establishing the statistical significance of the results obtained, and suitably visualizing the results of the project for consumption by relevant stakeholders.

5. This will be assessed through one of the assessment mechanisms in DSCI 345.

6. This will be assessed through one of the assessment mechanisms in PHIL 423.

7. This will be assessed through a detailed analysis of the DSCI 411 project reports, particularly the usefulness of the inferred information to the domain.

A full assessment plan will be devised to schedule these assessments over a 3-year cycle. The anticipated schedule will look as follows:

Year 1: Outcomes 1, 4, 5
Year 2: Outcomes 2, 4, 6
Year 3: Outcomes 3, 4, 7

c. Nature and level of research and/or scholarly work expected of program faculty; indicators of success in those areas.

Tenure-track faculty appointments will be made in various units throughout the university so as to maximize the interdisciplinary research strengths of the combined faculty. Such faculty will be expected to conduct innovative and impactful research, demonstrated by a history of successfully awarded grants, journal and or conference publications, and invited presentations.

Career NTTF appointments will be made in the Data Science Program, itself. There may also be TTF appointments in the program in the future if the research focus is on core data science research.

TTF and NTTF instructors are expected to deliver high-quality, engaging courses, as measured by student course evaluations, student success, and peer teaching evaluations.

6. Program Integration and Collaboration

a. Closely related programs in this or other Oregon colleges and universities.

OSU and OIT

b. Ways in which the program complements other similar programs in other Oregon institutions and other related programs at this institution. Proposal should identify the potential for collaboration.

Of the four major research universities in the state of Oregon (UO, OHSU, PSU and OSU), only OSU is proposing an undergraduate program with a domain emphasis structure that may be comparable to UO's. The OSU program in biological data science is a natural outgrowth of the Computational Genomics and Research Biotechnology (CGRB) program that has existed at OSU for many years, and with which there have been numerous connections with UO. The new OSU program has a primary focus on applying data science to biology. As such, it offers three options: computational biology, ecological and

environmental informatics, and genomics. Each of these options are focused in the natural sciences generally, and specifically in the life sciences.

In contrast, UO's program focuses first on data science, inclusive of disciplines across the university. This structure (so called data science + domain) is intended to contain domain emphases from across the schools and colleges at UO, and currently includes exemplar domain emphases in the natural sciences (biology), social sciences (geography), humanities (linguistics), and from the Lundquist College of Business (accounting analytics and marketing analytics). Therefore, the degree program at the University of Oregon is meant to be more broad than the program at OSU and cover all aspects of the social sciences, humanities, and professional schools, as well as the natural sciences.

c. If applicable, proposal should state why this program may not be collaborating with existing similar programs.

The other program in Oregon is at Oregon Institute of Technology (OIT). However, OIT's program is not comparable as it is not currently structured to have a domain emphasis.

d. Potential impacts on other programs.

The impacts on OSU and UO are likely to be mutually beneficial. The two institutions have complementary strengths in the area of biological data science (e.g., increased model organism research occurring at UO and increased environmental life sciences interests at OSU). Importantly, research and educational collaborations in these areas have occurred to varying degrees over the past decade.

Conversations between UO and OSU have been ongoing. For example, among Brett Tyler, director of CGRB at OSU, Bill Cresko in his role as executive director of DSI at UO, and Irem Tumer and David Conover, the vice presidents of research from OSU and UO, respectively. In addition, Director Cresko served on the external review committee for the CGRB at OSU at the request of Vice President Tumer, providing a deep understanding of the structure of the OSU program.

Institution: University of Oregon
Program: Bachelor of Arts and Bachelor of Science in Data Science

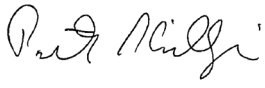
Action: At the **January 2, 2020** meeting, the Statewide Provosts Council approved a new program for **University of Oregon, BA/BS in Data Science,** to move forward to the Oregon Higher Education Coordinating Commission for its review and approval. The **University of Oregon** Board of Trustees has a tentative approval date of **January 27, 2020**.

Eastern Oregon University
Sarah Witte, provost
X Approved
__Opposed
__Abstained

Oregon State University
Ed Feser, provost
X Approved
__Opposed
__Abstained

Portland State University
Susan Jeffords, provost
X Approved
__Opposed
__Abstained

University of Oregon
Patrick Phillips, provost
X Approved
__Opposed
__Abstained



Oregon Health & Science University
Elena Andresen, interim provost
X Approved
__Opposed
__Abstained

Oregon Tech
Joanna Mott, provost
X Approved
__Opposed
__Abstained

Southern Oregon University
Susan Walsh, provost
X Approved
__Opposed
__Abstained

Western Oregon University
Rob Winningham, provost
X Approved
__Opposed
__Abstained